Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v10n2p349

A comprehensive study of lognormality tests

By Mahdizadeh, Zamanzade

Published: 14 October 2017

Electronic Journal of Applied Statistical Analysis
Vol. 10, Issue 02, October 2017, 349-373
DOI: 10.1285/i20705948v10n2p349

A comprehensive study of lognormality tests

M. Mahdizadeh*[a] and Ehsan Zamanzade[b]

[a]*Department of Statistics, Hakim Sabzevari University, P.O. Box 397, Sabzevar, Iran*
[b]*Department of Statistics, University of Isfahan, Isfahan, 81746-73441, Iran*

Published: 14 October 2017

There is abundant and increasing evidence that the lognormal distribution can account for random variation present in the data from many scientific fields. In the light of this flexibility for modeling, this article deals with goodness-of-fit tests for the lognormal distribution. Several testing procedures are compared by means of extensive simulation. Lastly, an actuarial data set is analyzed for illustration.

Keywords: Empirical distribution function, Information theory, Likelihood ratio.

1 Introduction

The Gaussian distribution is most often assumed to describe the random variation occurring in the data from many scientific disciplines. Many measurements, however, reveal a more or less skewed distribution. The lognormal (LN) distribution can be a potential model if data are non-negative and right skewed, while log-transformed data show normality properties. The corresponding probability density function (PDF) is given by

$$f_0(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma x}\exp\left\{-\frac{(\log(x)-\mu)^2}{2\sigma^2}\right\}, \quad \mu\in\mathbb{R}, \sigma>0; x>0 \qquad (1)$$

which is simply denoted by $X \sim LN(\mu,\sigma^2)$.

In the sequel, we provide some applications of the distribution in different fields. Koch (1966, 1969) has discussed the genesis of the LN distribution arising from biological

*Corresponding author: mahdizadeh.m@live.com.

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

and pharmacological mechanisms. For example, he considered the LN distribution for modeling the metabolic turnover. Applications of this distribution in biochemistry are given in Masuyama (1984) and references therein. It has also been utilized in medical studies such as the incubation period of disease, the time to recovery, and duration of survival (Crow and Shimizu, 1988, pp. 211-266). In the majority of plant and animal communities, the abundance of species follows a (truncated) LN distribution (Magurran, 1988). Hydrological variables, like low river flows and water quality data, are also often modelled as LN random variables (Kroll and Stedinger, 1996). In the actuarial context, models with heavy-tailed distributions have been used to provide adequate descriptions of claim size distributions. The right tail of a distribution is an important issue in insurance as it represents the total impact of insurance losses. The LN distribution is a flexible choice which can be used in modeling claim size (Boland, 2007).

The use of a parametric model needs to be supported by a formal testing procedure based on the observed data. Otherwise, inferential results under the hypothesized model could be quite misleading. Thanks to importance of the LN model for data analysis in many scientific disciplines (as exemplified earlier), goodness-of-fit tests for the LN distribution deserve more attention. This work provides a comparative power study of such tests through extensive simulation.

Section 2 presents a review of existing tests. Section 3 contains results of a Monte Carlo study conducted to compare power of the tests. A data set from the actuarial literature is analyzed in Section 4. We end in Section 5 with a summary. Figures are collected in an appendix.

2 Goodness-of-fit tests

Given a random sample $X_1, \ldots, X_n$ from a population having a continuous density function $f(x)$, consider the problem of testing $H_0 : f(x) = f_0(x; \mu, \sigma^2)$ for some $\mu \in \mathbb{R}$ and $\sigma > 0$, where f_0 is defined in (1). The alternative hypothesis is $H_1 : f(x) \neq f_0(x; \mu, \sigma^2)$ for any $\mu \in \mathbb{R}$ and $\sigma > 0$.

2.1 Tests based on transformed data

Owing to the definition of the LN distribution, the above problem can be reduced to that of testing normality for the log transformed data. Many statistical procedures rely on the assumption that the observed data are normally distributed. Consequently, there exists a vast literature on tests of normality and their statistical properties. In this subsection, it is assumed that $\mathcal{X}_i = \log(X_i)$ for $i = 1, \ldots, n$.

Pearson's chi-squared test can be used to determine whether the population distribution estimated by a random sample containing n independent observations is identical to some hypothesized distribution (Kirk, 2008). The test statistic is given by

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}, \tag{2}$$

where O_i's are the observed frequencies in k mutually exclusive categories of a population, and E_i's are the expected frequencies under the null hypothesis. The null distribution of χ^2 is approximately chi-squared with $k-c-1$ degrees of freedom, where c is the number of estimated parameters for the distribution to be tested. These parameters are used in computing E_i's.

It is well known that the most powerful omnibus test of normality in the literature is the Shapiro-Wilk (SW) which is essentially the squared ratio of the best linear unbiased estimator for scale to the standard deviation. The test statistic is defined by

$$\mathrm{SW} = \frac{\left(\sum_{i=1}^{n} a_i \mathcal{X}_{(i)}\right)^2}{\sum_{i=1}^{n}\left(\mathcal{X}_i - \bar{\mathcal{X}}\right)^2}, \tag{3}$$

where $\mathcal{X}_{(i)}$'s are the sample order statistics and $\bar{\mathcal{X}}$ is the sample mean. The vector $\mathbf{a}' = (a_1, \ldots, a_n)$ is given by

$$\mathbf{a} = \frac{\mathbf{m}'\mathbf{V}^{-1}}{\left(\mathbf{m}'\mathbf{V}^{-1}\mathbf{V}^{-1}\mathbf{m}\right)^{1/2}},$$

where $\mathbf{m}$ is the vector of the expected values of the order statistics of a simple random sample of size n from the standard normal distribution, and $\mathbf{V}$ is the covariance matrix of those order statistics. For $n \leq 50$, the values of a_i were tabulated by Shapiro and Wilk (1965).

The Shapiro-Francia (SF) test was described in Shapiro and Francia (1972). It is obtained through replacing V^{-1} in the SW test by the identity matrix I, when the sample size is large. The statistic is then

$$\mathrm{SF} = \frac{\left(\sum_{i=1}^{n} b_i \mathcal{X}_{(i)}\right)^2}{\sum_{i=1}^{n}\left(\mathcal{X}_i - \bar{\mathcal{X}}\right)^2}, \tag{4}$$

where the vector $\mathbf{b}' = (b_1, \ldots, b_n)$ is given by

$$\mathbf{b} = \frac{\mathbf{m}'}{\left(\mathbf{m}'\mathbf{m}\right)^{1/2}}.$$

Royston (1993a, 1993b) proposed an easy-to-calculate approximation to the SF test and its p-value for sample sizes $5 \leq n \leq 5000$.

Jarque-Bera (JB) test, introduced by Jarque and Bera (1987), is based on a comparison of the standardized sample skewness and kurtosis to the skewness and kurtosis of the standard normal distribution, which are 0 and 3, respectively. The form of the statistic is given by

$$\mathrm{JB} = n\left(\frac{\sqrt{g_1}^{\,2}}{6} + \frac{(g_2-3)^2}{24}\right), \tag{5}$$

where $\sqrt{g_1} = m_3/m_2^{3/2}$ and $g_2 = m_4/m_2^2$, with $m_j = \sum_{i=1}^{n}(\mathcal{X}_i - \bar{\mathcal{X}})^j/n$. The JB statistic asymptotically follows chi-squared distribution with two degrees of freedom.

Some tests employ measures of discrepancy between the empirical distribution function (EDF) and a given distribution function. There are two classes of such tests: the Kolmogorov-Smirnov type and the quadratic type. We now present some EDF-based tests of normality. To do so, we assume that Z_i's are the standardized data,

$$Z_i = \frac{\mathcal{X}_i - \bar{\mathcal{X}}}{\sqrt{\sum_{j=1}^{n}(\mathcal{X}_j - \bar{\mathcal{X}})^2/n}} \quad (i = 1, \ldots, n),$$

with ordered values $Z_{(i)}$'s. Also, the cumulative distribution function of a standard normal random variable is denoted by $\Phi(.)$.

The best-known statistic for tests of fit is that of Kolmogorov-Smirnov given by

$$D = \max_{i=1,\ldots,n}\left[\max\left\{\frac{i}{n} - \Phi(Z_{(i)}), \Phi(Z_{(i)}) - \frac{i-1}{n}\right\}\right]. \tag{6}$$

The famous Cramér-von Mises statistic,

$$W^2 = \sum_{i=1}^{n}\left(\Phi(Z_{(i)}) - \frac{i-0.5}{n}\right)^2 + \frac{1}{12n}, \tag{7}$$

leads to an important goodness-of-fit test. Another powerful test, especially for small sample sizes, is based on the Anderson-Darling statistic defined as

$$A^2 = -\frac{2}{n}\sum_{i=1}^{n}\left[(i-0.5)\log\left\{\Phi(Z_{(i)})\right\} + (n-i+0.5)\log\left\{1-\Phi(Z_{(i)})\right\}\right] - n. \tag{8}$$

Details about the derivation of the above three statistics are given in Gibbons and Chakraborti (2011). Kuiper (1960) introduced a statistic related to that of Kolmogorov-Smirnov to be applied to observations on a circle. The statistic has the form

$$V = \max_{i=1,\ldots,n}\left\{\frac{i}{n} - \Phi(Z_{(i)})\right\} + \max_{i=1,\ldots,n}\left\{\Phi(Z_{(i)}) - \frac{i-1}{n}\right\}. \tag{9}$$

Watson (1961) suggested a modification of (7) given by

$$U^2 = W^2 - n\left(\frac{1}{n}\sum_{i=1}^{n}\Phi(Z_{(i)}) - 0.5\right)^2. \tag{10}$$

Information theoretic measures have been widely used in developing tests of fit for different parametric families. A classical measure of uncertainty for a PDF $f(x)$ is the differential entropy, also known as the Shannon information measure, defined as

$$H(f) = -\int_{-\infty}^{\infty} f(x)\log f(x)\mathrm{d}x.$$

The normal distribution possesses the largest entropy among all distributions with the same variance (Shannon and Weaver, 1949). One can build on this result to construct a test of normality. Vasicek (1976) introduced a simple entropy estimator given by

$$HV_{m,n} = \frac{1}{n}\sum_{i=1}^{n} \log\left\{\frac{n}{2m}\left(\mathcal{X}_{(i+m)} - \mathcal{X}_{(i-m)}\right)\right\}, \tag{11}$$

where m (called window size) is a positive integer less than or equal to $n/2$, $\mathcal{X}_{(1)} \leq \cdots \leq \mathcal{X}_{(n)}$ are order statistics based on a random sample of size n, $\mathcal{X}_{(i)} = \mathcal{X}_{(1)}$ for $i < 1$, and $\mathcal{X}_{(i)} = \mathcal{X}_{(n)}$ for $i > n$. He showed that (11) is a consistent estimator of the population entropy. In particular, $HV_{m,n} \xrightarrow{p} H(f)$ as $m \to \infty$, $n \to \infty$ and $m/n \to 0$, where $\xrightarrow{p}$ denotes convergence in probability. Vasicek (1976) suggested using the following statistic for testing normality

$$T_V = \frac{\exp\{HV_{m,n}\}}{\sqrt{\sum_{i=1}^{n}(\mathcal{X}_i - \bar{\mathcal{X}})^2/n}}. \tag{12}$$

It is possible to modify (12) by using improved entropy estimators motivated by Vasicek's estimator. Some of these estimators are set out here. Bowman (1992) studied the estimator

$$HB_n = -\frac{1}{n}\sum_{i=1}^{n} \log\left\{\hat{f}(\mathcal{X}_i)\right\}, \tag{13}$$

where

$$\hat{f}\,(x) = \frac{1}{nh}\sum_{j=1}^{n} K\left(\frac{x - \mathcal{X}_j}{h}\right),$$

and $K(.)$ is a symmetric kernel function which is chosen to be the standard normal density function. The bandwidth h is selected based on the normal optimal smoothing formula, $h = 1.06\, s\, n^{-1/5}$, where s is the sample standard deviation.

Van Es (1992) considered estimation of functionals of a PDF and entropy in particular. He proposed the following estimator

$$HVE_{m,n} = \frac{1}{n-m}\sum_{i=1}^{n-m} \log\left\{\frac{n+1}{m}\left(\mathcal{X}_{(i+m)} - \mathcal{X}_{(i)}\right)\right\} + \sum_{i=m}^{n}\frac{1}{i} - \log\left\{\frac{n+1}{m}\right\}, \tag{14}$$

where m is a positive integer less than n.

Correa (1995) proposed another entropy estimator defined as

$$HC_{m,n} = -\frac{1}{n}\sum_{i=1}^{n} \log\left\{\frac{\sum_{j=i-m}^{i+m}\left(\mathcal{X}_{(j)} - \bar{\mathcal{X}}_{(i)}\right)(j-i)}{n\sum_{j=i-m}^{i+m}\left(\mathcal{X}_{(j)} - \bar{\mathcal{X}}_{(i)}\right)^2}\right\}, \tag{15}$$

where

$$\bar{\mathcal{X}}_{(i)} = \frac{1}{2m+1}\sum_{j=i-m}^{i+m} \mathcal{X}_{(j)}.$$

Zamanzadeh and Arghami (2009) modified (15) by assigning different weights at the boundaries. Their estimator has the form

$$HZ_{m,n} = -\frac{1}{n}\sum_{i=1}^{n}\log\left\{\frac{\sum_{j=k_1(i)}^{k_2(i)}\left(\mathcal{X}_{(j)}-\tilde{\mathcal{X}}_{(i)}\right)\left(j-\tilde{F}_i\right)}{n\sum_{j=k_1(i)}^{k_2(i)}\left(\mathcal{X}_{(j)}-\tilde{\mathcal{X}}_{(i)}\right)^2}\right\}, \tag{16}$$

where

$$\tilde{\mathcal{X}}_{(i)} = \frac{1}{k_2(i)-k_1(i)+1}\sum_{j=k_1(i)}^{k_2(i)}\mathcal{X}_{(j)},$$

$$\tilde{F}_i = \frac{1}{k_2(i)-k_1(i)+1}\sum_{j=k_1(i)}^{k_2(i)} j,$$

$k_1(i) = \max\{1, i-m\}$ and $k_2(i) = \min\{n, i+m\}$.

Yousefzadeh and Arghami (2008) introduced the following entropy estimator

$$HY_{m,n} = \sum_{i=1}^{n}\left\{\frac{\hat{F}_y\left(\mathcal{X}_{(i+m)}\right)-\hat{F}_y\left(\mathcal{X}_{(i-m)}\right)}{\sum_{j=1}^{n}\hat{F}_y\left(\mathcal{X}_{(j+m)}\right)-\hat{F}_y\left(\mathcal{X}_{(j-m)}\right)}\right\}\log\left\{\frac{\mathcal{X}_{(i+m)}-\mathcal{X}_{(i-m)}}{\hat{F}_y\left(\mathcal{X}_{(i+m)}\right)-\hat{F}_y\left(\mathcal{X}_{(i-m)}\right)}\right\}, \tag{17}$$

where for $i = 2, \ldots, n-1$,

$$\hat{F}_y\left(\mathcal{X}_{(i)}\right) = \frac{n-1}{n(n+1)}\left(i+\frac{1}{n-1}+\frac{\mathcal{X}_{(i)}-\mathcal{X}_{(i-1)}}{\mathcal{X}_{(i+1)}-\mathcal{X}_{(i-1)}}\right),$$

and

$$\hat{F}_y\left(\mathcal{X}_{(1)}\right) = 1-\hat{F}_y\left(\mathcal{X}_{(n)}\right) = \frac{1}{n+1}.$$

The test statistics obtained by replacing $HV_{m,n}$ in (12) with HB_n, $HVE_{m,n}$, $HC_{m,n}$, $HY_{m,n}$ and $HZ_{m,n}$ will be denoted by T_B, T_{VE}, T_C, T_Y and T_Z, respectively.

Zhang and Wu (2005) developed three tests of normality based on the likelihood ratio statistic. The corresponding test statistics are

$$Z_K = \max_{i=1,\ldots,n}\left[(i-0.5)\log\left\{\frac{i-0.5}{n\Phi(Z_{(i)})}\right\}+(n-i+0.5)\log\left\{\frac{n-i+0.5}{n(1-\Phi(Z_{(i)}))}\right\}\right], \tag{18}$$

$$Z_A = -\sum_{i=1}^{n}\left[\frac{\log\left\{\Phi(Z_{(i)})\right\}}{n-i+0.5}+\frac{\log\left\{1-\Phi(Z_{(i)})\right\}}{i-0.5}\right], \tag{19}$$

and

$$Z_C = \sum_{i=1}^{n}\left[\log\left\{\frac{1/\Phi(Z_{(i)})-1}{(n-0.5)/(i-0.75)-1}\right\}\right]^2. \tag{20}$$

2.2 Tests based on original data

To the best of our knowledge, there are few testing procedures employing the original data. Recently, Batsidis et al. (2016) developed two testing procedures based on the Kullback-Leibler (KL) distance which is an extended concept of the entropy. The KL distance between two PDFs f and g is defined to be

$$I(f,g)=\int_{-\infty}^{\infty} f(x)\log\left(\frac{f(x)}{g(x)}\right)\mathrm{d}x.$$

It is well known that $I(f,g)\geq 0$ and the equality holds if and only if $f(x)=g(x)$, almost surely.

Let X be a nonnegative random variable with PDF f such that $\mu_r=E(X^r)<\infty$. Then, the r-size weighted distribution f_r exists and its PDF is given by

$$f_r(x)=\frac{x^r}{\mu_r}f(x).$$

Tzavelas and Economou (2012) established the following characterization of the LN distribution.

Proposition 1. *If $X\sim LN(\mu,\sigma^2)$, then $I(f,f_r)=I(f_r,f)$, $\forall r\in\mathbb{R}$. Conversely, if $I(f,f_r)=I(f_r,f)$, $\forall r\in\mathbb{O}$ for some interval $\mathbb{O}$ containing 0, then $X\sim LN(\mu,\sigma^2)$.*

Proposition 1 can be equivalently expressed in terms of

$$\delta_r=I(f,f_r)-I(f_r,f)=2\log\left(E(X^r)\right)-E\left(\log(X^r)\right)-r\frac{E\left(X^r\log(X)\right)}{E(X^r)}.$$

Put it another way, if $X\sim LN(\mu,\sigma^2)$, then $\delta_r=0$, $\forall r\in\mathbb{R}$, while if $\delta_r=0$, $\forall r\in\mathbb{O}$ for some interval $\mathbb{O}$ containing 0, then $X\sim LN(\mu,\sigma^2)$. A test for lognormality therefore can be based on deviation of an empirical estimator of δ_r from zero.

Suppose $\hat{\sigma}^2$ is the maximum likelihood estimator of the shape parameter σ^2 in (1) given by

$$\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}\left(\log(X_i)-\hat{\mu}\right)^2,$$

where $\hat{\mu}=\sum_{i=1}^{n}\log(X_i)/n$. Based on the transformed data $\hat{Y}_i=X_i^{1/\hat{\sigma}}$, an estimator of δ_r can be made as

$$D_{n,r}=2\log\left(\frac{1}{n}\sum_{i=1}^{n}\hat{Y}_i^{\,r}\right)-\frac{r}{n}\sum_{i=1}^{n}\log(\hat{Y}_i)-r\frac{\sum_{i=1}^{n}\hat{Y}_i^{\,r}\log(\hat{Y}_i)}{\sum_{i=1}^{n}\hat{Y}_i^{\,r}}. \quad (21)$$

To set forth the tests based on $D_{n,r}$, we borrow two propositions from Batsidis et al. (2016). The first one deals with the asymptotic null distribution of $D_{n,r}$ for a fixed r.

Proposition 2. *Let $X_1,\ldots,X_n$ be a random sample from $LN(\mu,\sigma^2)$, and $D_{n,r}$ be as in (21). Then $\sqrt{n}D_{n,r}\xrightarrow{d}N(0,\sigma_r^2)$, where $\xrightarrow{d}$ denotes convergence in distribution, and*

$$\sigma_r^2=\exp(r^2)(r^4-3r^2+4)-(r^2+4). \quad (22)$$

Based on Proposition 2, one can test the hypothesis of lognormality by evaluating $D_{n,r}$ at a single point $r \neq 0$, or at a finite grid of points. Following this approach, we may lose the information given by using the rest of the points. In addition, seeking values of the argument r which give rise to good power trait is not an easy task. To sidestep these problems, a different approach needs to be pursued. The next proposition attends to behavior of $D_{n,r}/\sigma_r$, and paves the way for introducing the test statistics.

Proposition 3. *Let* $X_1, \ldots, X_n$ be a random sample from $LN(\mu, \sigma^2)$, and $D_{n,r}$ and σ_r^2 be defined as in (21) and (22), respectively. It holds that

$$\sqrt{n} \sup_{r \in \mathbb{R}} \left| \frac{D_{n,r}}{\sigma_r} \right|$$

and

$$n \int_{-\infty}^{\infty} \left(\frac{D_{n,r}}{\sigma_r} \right)^2 \mathrm{d}r$$

are bounded in probability.

In view of the above result, the statistics

$$T_1 = \sqrt{n} \sup_{r \in \mathbb{R}} \left| \frac{D_{n,r}}{\sigma_r} \right|$$

and

$$T_2 = n \int_{-\infty}^{\infty} \left(\frac{D_{n,r}}{\sigma_r} \right)^2 \mathrm{d}r$$

are well defined and can be used to test the hypothesis of lognormality. In practice, T_1 and T_2 are approximated by

$$T_1 \approx \sqrt{n} \sup_{r \in (-5,5)} \left| \frac{D_{n,r}}{\sigma_r} \right|, \tag{23}$$

and

$$T_2 \approx n \int_{-5}^{5} \left(\frac{D_{n,r}}{\sigma_r} \right)^2 \mathrm{d}r, \tag{24}$$

respectively. As mentioned by Batsidis et al. (2016), these forms lead to a significant simplification of the test statistics.

Remark 1: The family of LN distribution is invariant under the transformation $\exp(a)X^b$, for $a, b \in \mathbb{R}$. Putting this and Proposition 2 together, it follows that for fixed r, the limit distributions of T_1 and T_2 under the hypothesis of lognormality is independent of μ and σ^2. Hence, study under the null hypothesis can be restricted to the case of LN(0,1).

2.3 Determination of critical values

It is generally difficult to derive exact distributions of the test statistics presented in the previous subsections under the null hypothesis although asymptotic approximations are available in some cases. Monte Carlo simulation was then employed to determine critical

Table 1: The optimal window sizes for the tests of size 0.05 based on the KL distance.

	Statistic				
n	T_V	T_{VE}	T_C	T_Y	T_Z
10	3	8	4	5	5
20	3	19	3	10	4
30	4	29	4	15	4
50	5	49	4	25	5

values of a generic test statistic, say T, whose large values imply rejection of H_0. To this end, 20,000 samples of size n were generated from $N(0,1)$ or $LN(0,1)$ (depending on the fact that T is based on transformed data or original data), and the statistic was computed from each sample. Finally, $(1-\alpha)$ quantile of the resulting values was determined which will be denoted by $\mathcal{T}_{n,1-\alpha}$. The composite null hypothesis is rejected at level α if the observed value of T exceeds $\mathcal{T}_{n,1-\alpha}$.

Remark 2: If small values of T support rejection of H_0 (this is the case for SW, SF and the entropy based tests), then $\mathcal{T}_{n,\alpha}$, α quantile of 20,000 values of T, would be the appropriate critical value.

To calculate test statistics based on the entropy (with the exception of T_B), the window size m corresponding to a given sample size must be selected in advance. In entropy estimation based on spacings, choosing optimal m for given n is still an open problem. For each n, the window size having largest critical value tends to yield greater power. For sample sizes 10, 20, 30 and 50, window sizes producing the maximum critical values for different tests are given in Table 1.

Remark 3: In the all entropy estimators which employ spacings of the order statistics, it is assumed that m is an integer satisfying $1 \le m \le n/2$, unless otherwise stated.

Table 2 contains 0.05 critical points for most of the tests considered in this study. For the entropy based tests, the above mentioned optimal window sizes are used. These thresholds will be used in the next section to study the power properties. If a test statistic is not included in Table 2, it means that the p-value needed to perform the test is widely accessible through statistical softwares.

3 Numerical results

In this section, performances of the proposed tests are evaluated via Monte Carlo experiments. Toward this end, we considered seven families of alternatives:

- The LN distribution with PDF

$$f(x) = \frac{1}{\sqrt{2\pi}\theta x} \exp\left\{-\frac{(\log x)^2}{2\theta^2}\right\}, \quad \theta > 0; x > 0,$$

Table 2: 0.05 critical points of the tests.

	n			
Statistic	10	20	30	50
D	0.265	0.193	0.160	0.125
W^2	0.124	0.125	0.127	0.126
A^2	0.725	0.743	0.765	0.754
V	0.435	0.317	0.263	0.205
U^2	0.115	0.116	0.118	0.116
T_V	2.195	2.770	3.065	3.361
T_B	3.718	3.775	3.825	3.899
T_{VE}	3.491	3.616	3.783	4.022
T_C	2.694	3.166	3.429	3.680
T_Y	3.541	3.593	3.674	3.809
T_Z	3.399	3.469	3.660	3.839
Z_K	1.048	1.382	1.618	1.861
Z_A	3.502	3.451	3.423	3.387
Z_C	6.695	9.160	10.849	12.643
T_1	1.555	1.854	1.955	2.081
T_2	2.697	4.015	4.603	5.690

denoted by $LN(\theta)$.

- The gamma distribution with PDF

$$f(x) = \frac{1}{\Gamma(\theta)} x^{\theta-1} \exp(-x), \quad \theta > 0; x > 0,$$

denoted by $G(\theta)$.

- The folded normal distribution defined as $|Z|$ with $Z \sim N(0,1)$, and denoted by FN.

- The beta distribution with PDF

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad \alpha, \beta > 0; 0 < x < 1,$$

denoted by $B(\alpha, \beta)$.

- The shifted Pareto distribution with PDF

$$f(x) = \theta(1+x)^{-(\theta+1)}, \quad \theta > 0; x > 0,$$

denoted by $Pa(\theta)$.

- The generalized exponential distribution with PDF

$$f(x) = \theta e^{-x} \left(1 - e^{-x}\right)^{\theta-1}, \quad \theta > 0; x > 0,$$

denoted by $GE(\theta)$.

- Johnson's distribution defined as

$$\frac{\exp\left(\frac{Z-\theta}{\theta}\right)}{1+\exp\left(\frac{Z-\theta}{\theta}\right)}$$

with $Z \sim N(0,1)$, and denoted by $Sb(\theta)$.

The members selected from the above families are LN(0.5), LN(1), LN(2), G(1), G(1.5), G(2), FN, B(0.5,2), Pa(2), Pa(4), Pa(6), GE(0.5), GE(2), GE(3), Sb(0.5) and Sb(1). For each alternative, 20,000 samples of sizes $n = 10, 30$ were generated, and the power of each test was estimated by the percentages of samples entering the rejection region. Tables 3-6 present the estimated powers of the twenty tests of size 0.05 described in the previous section. Given a sample size and alternative, power entry associated with the best test is in bold.

It is observed that all tests maintain the nominal level satisfactorily. So we may fairly judge the tests based on their powers. Increasing the sample size usually results in power improvement for each test. For a fixed alternative, power of the tests are not markedly different when $n = 10$, while this is not the case for $n = 30$.

Table 3: Power comparison for the tests of size 0.05 for $n = 10$.

	Alternative							
Statistic	LN(0.5)	LN(1)	LN(2)	G(1)	G(1.5)	G(2)	FN	B(0.5,2)
χ^2	0.066	0.068	0.067	0.122	0.100	0.096	0.186	0.269
D	0.046	0.053	0.053	0.114	0.092	0.084	0.170	0.223
W^2	0.047	0.048	0.050	0.123	0.099	0.087	0.194	0.271
A^2	0.049	0.047	0.050	0.132	0.107	0.094	0.212	0.296
V	0.049	0.049	0.050	0.102	0.084	0.077	0.161	0.230
U^2	0.047	0.048	0.051	0.113	0.092	0.082	0.179	0.249
SW	0.049	0.047	0.052	0.146	0.119	0.101	0.232	0.322
SF	0.053	0.051	0.054	**0.157**	**0.129**	**0.110**	0.243	0.321
JB	0.049	0.047	0.049	0.149	0.123	0.105	0.219	0.271
T_V	0.048	0.047	0.049	0.104	0.086	0.076	0.176	0.286
T_B	0.053	0.052	0.054	0.157	0.128	0.109	0.233	0.304
T_{VE}	0.050	0.046	0.050	0.122	0.103	0.087	0.176	0.233
T_C	0.048	0.046	0.052	0.114	0.092	0.083	0.191	0.306
T_Y	0.048	0.048	0.049	0.105	0.085	0.076	0.170	0.275
T_Z	0.050	0.051	0.053	0.144	0.114	0.097	0.228	0.323
Z_K	0.046	0.051	0.051	0.108	0.086	0.082	0.163	0.219
Z_A	0.050	0.048	0.052	0.155	0.126	0.107	**0.244**	**0.332**
Z_C	0.052	0.048	0.053	0.152	0.125	0.107	0.241	0.329
T_1	0.049	0.047	0.049	0.150	0.123	0.105	0.222	0.276
T_2	0.048	0.046	0.048	0.149	0.123	0.105	0.222	0.278

Table 4: Power comparison for the tests of size 0.05 for $n = 10$.

	Alternative							
Statistic	Pa(2)	Pa(4)	Pa(6)	GE(0.5)	GE(2)	GE(3)	Sb(0.5)	Sb(1)
χ^2	0.086	0.105	0.109	0.192	0.088	0.077	0.245	0.097
D	0.075	0.095	0.098	0.181	0.075	0.066	0.182	0.079
W^2	0.076	0.099	0.106	0.209	0.078	0.067	0.217	0.084
A^2	0.081	0.106	0.115	0.226	0.084	0.071	0.238	0.089
V	0.069	0.086	0.090	0.171	0.069	0.064	0.194	0.077
U^2	0.072	0.092	0.099	0.192	0.073	0.064	0.205	0.080
SW	0.088	0.114	0.124	0.249	0.090	0.075	0.261	0.094
SF	**0.103**	**0.125**	**0.135**	0.257	**0.101**	**0.081**	0.241	0.099
JB	0.101	0.121	0.129	0.231	0.097	0.075	0.167	0.087
T_V	0.061	0.081	0.088	0.188	0.067	0.060	0.288	0.087
T_B	0.101	0.124	0.133	0.248	0.100	0.080	0.212	0.094
T_{VE}	0.083	0.100	0.107	0.187	0.081	0.066	0.189	0.084
T_C	0.067	0.086	0.096	0.202	0.074	0.063	**0.289**	0.090
T_Y	0.062	0.081	0.088	0.186	0.069	0.060	0.263	0.087
T_Z	0.083	0.109	0.118	0.240	0.087	0.073	0.249	0.099
Z_K	0.074	0.090	0.096	0.174	0.073	0.064	0.184	0.080
Z_A	0.099	0.122	0.132	**0.259**	0.098	0.078	0.258	0.098
Z_C	0.096	0.121	0.130	0.256	0.098	0.078	0.261	**0.100**
T_1	0.101	0.121	0.129	0.233	0.096	0.075	0.172	0.088
T_2	0.100	0.121	0.129	0.234	0.096	0.075	0.174	0.088

Table 5: Power comparison for the tests of size 0.05 for $n = 30$.

	Alternative							
Statistic	LN(0.5)	LN(1)	LN(2)	G(1)	G(1.5)	G(2)	FN	B(0.5,2)
χ^2	0.053	0.052	0.051	0.185	0.134	0.108	0.346	0.807
D	0.050	0.045	0.046	0.280	0.201	0.161	0.453	0.716
W^2	0.046	0.046	0.046	0.344	0.247	0.190	0.560	0.858
A^2	0.043	0.044	0.043	0.374	0.266	0.206	0.600	0.904
V	0.050	0.048	0.047	0.254	0.180	0.141	0.448	0.831
U^2	0.047	0.047	0.046	0.295	0.211	0.163	0.500	0.814
SW	0.050	0.050	0.048	0.462	**0.339**	0.272	0.693	0.955
SF	0.052	0.051	0.050	0.454	0.337	0.273	0.674	0.929
JB	0.050	0.046	0.051	0.413	0.313	0.254	0.593	0.758
T_V	0.048	0.048	0.047	0.301	0.204	0.155	0.546	0.964
T_B	0.046	0.045	0.045	0.406	0.298	0.236	0.618	0.872
T_{VE}	0.054	0.050	0.052	0.047	0.050	0.047	0.055	0.186
T_C	0.051	0.050	0.048	0.289	0.195	0.147	0.531	0.958
T_Y	0.053	0.048	0.047	0.110	0.087	0.074	0.217	0.715
T_Z	0.050	0.051	0.048	0.299	0.202	0.152	0.544	0.956
Z_K	0.044	0.045	0.042	0.342	0.246	0.195	0.550	0.933
Z_A	0.046	0.046	0.046	**0.465**	**0.339**	0.272	**0.704**	**0.971**
Z_C	0.047	0.047	0.046	0.451	0.331	0.266	0.677	0.949
T_1	0.054	0.051	0.055	0.448	0.338	0.279	0.637	0.805
T_2	0.052	0.049	0.055	0.449	**0.339**	**0.280**	0.638	0.802

Table 6: Power comparison for the tests of size 0.05 for $n = 30$.

	Alternative							
Statistic	Pa(2)	Pa(4)	Pa(6)	GE(0.5)	GE(2)	GE(3)	Sb(0.5)	Sb(1)
χ^2	0.086	0.116	0.136	0.361	0.096	0.070	0.571	0.113
D	0.120	0.183	0.205	0.468	0.140	0.094	0.485	0.160
W^2	0.142	0.215	0.249	0.577	0.164	0.107	0.640	0.195
A^2	0.152	0.234	0.272	0.617	0.177	0.113	0.716	0.217
V	0.114	0.162	0.184	0.466	0.120	0.083	0.608	0.142
U^2	0.126	0.187	0.214	0.511	0.140	0.092	0.588	0.169
SW	0.201	0.299	0.345	0.708	0.234	0.155	0.832	0.290
SF	0.222	0.306	0.347	0.687	**0.236**	0.159	0.757	0.256
JB	0.224	0.295	0.331	0.609	0.227	0.156	0.459	0.197
T_V	0.086	0.160	0.197	0.572	0.135	0.086	**0.888**	0.242
T_B	0.199	0.272	0.311	0.631	0.205	0.135	0.641	0.213
T_{VE}	0.029	0.035	0.038	0.057	0.050	0.047	0.279	0.117
T_C	0.078	0.151	0.184	0.558	0.127	0.085	0.883	0.246
T_Y	0.039	0.061	0.071	0.232	0.068	0.056	0.611	0.158
T_Z	0.081	0.155	0.194	0.571	0.132	0.086	0.879	0.248
Z_K	0.152	0.222	0.255	0.572	0.168	0.113	0.788	0.193
Z_A	0.197	0.297	0.342	**0.719**	0.234	0.152	0.880	**0.303**
Z_C	0.203	0.294	0.340	0.693	0.231	0.153	0.818	0.276
T_1	**0.234**	0.314	**0.358**	0.648	0.248	0.171	0.522	0.228
T_2	0.230	**0.315**	**0.358**	0.649	0.249	**0.172**	0.525	0.232

Table 7: Indexed hurricane losses, 1949-1980 (000 omitted).

6766	16983	30146	49397	102942	198446	513586
7123	18383	33727	52600	103217	227338	545778
10562	19030	40596	59917	123680	329511	750389
14474	25304	41409	63123	140136	361200	863881
15351	29112	47905	77809	192013	421680	1638000

Although no single test is uniformly superior, the following conclusions can be made. Among the EDF-based tests, A^2 often shows the best performance. Within the class of entropy based tests, T_B is generally powerful against all distributions. Among the three tests based on the likelihood ratio statistic, Z_A commonly outperforms the others. Also, it emerges that result from T_1 and T_2 are in good agreement. Overall, SW and Z_A tests are close competitors that nearly achieve the highest powers (compare with bold entries), given a sample size and alternative.

Figures 1-5 facilitate power comparisons among the tests for $n = 10, 20, 30, 50$ under some alternative distributions. To have simpler plots, we only considered four tests A^2, SW, Z_A and T_1 which revealed good power properties. It is evident that SW and Z_A tests are preferable in many cases.

4 Application

In actuarial sciences, information about the process producing the losses is essential for setting premiums, evaluating the effects of deductibles and limits, and determining the impact of inflation. The LN distribution is an adaptable choice for modeling loss distributions. In this section, we analyze a data set of losses from hurricanes occurring from 1949 to 1980 as provided by the American Insurance Association. It was also applied by Hogg and Klugman (1983). Table 7 contains the losses, where all values are in 1981 dollars and only those greater than 5,000,000 have been included.

It is of interest to check whether the LN model could fit the data. If we find some positive evidence, then the distribution can be used to estimate the frequency of losses in excess of a specified amount, or to estimate the expected number of years between hurricanes causing a specified amount of damage. The LN Q-Q plot appears in Figure 6. The histogram along with the corresponding fitted LN curve is also included. The maximum likelihood estimates of the parameters are $\hat{\mu} = 11.2411$ and $\hat{\sigma} = 1.4271$.

The values of all statistics are reported in Table 8, where the numbers in parenthesis indicate the associated 0.05 critical points. As to χ^2, SW, SF and JB tests, the numbers in parenthesis show the corresponding p-values. The hypothesis testing is done by comparing the value of each statistic with the associated critical point, and considering whether small or large values of the statistic support rejection of H_0. Also, if the significance level is larger than the p-value corresponding to a statistic, then H_0 is rejected.

Table 8: Observed values of the different statistics, where the numbers in parenthesis indicate either the the associated 0.05 critical points, or p-values.

χ^2	D	W^2	A^2	V	U^2	SW	SF	JB	T_V
2.800	0.095	0.054	0.328	0.171	0.050	0.972	0.980	1.470	3.410
(0.833)	(0.148)	(0.125)	(0.747)	(0.245)	(0.116)	(0.499)	(0.673)	(0.304)	(3.159)
T_B	T_{VE}	T_C	T_Y	T_Z	Z_K	Z_A	Z_C	T_1	T_2
4.066	4.076	3.692	3.864	3.822	0.515	3.323	5.666	0.906	1.034
(3.855)	(3.849)	(3.511)	(3.708)	(3.720)	(1.665)	(3.408)	(11.144)	(1.993)	(4.943)

By using any test, the null hypothesis that the data follow the LN distribution is not rejected at 0.05 significance level.

5 Conclusion

The LN distribution enjoys desirable flexibility in modeling a variety of phenomena. It has found applications in many fields including biology, earth sciences, ecology, economics, insurance, medicine and software reliability engineering, among others. Statistical inference in parametric setup is prone to violation of distributional assumption. This, in turn, necessitates developing formal testing procedures for the LN model.

This article aims to provide an exhaustive study of goodness-of-fit tests for the LN distribution. Twenty tests available in the literature were selected which includes EDF-based tests, information theoretic tests, tests based on regression or sample moments, and some tests derived from the likelihood ratio statistic. Extensive simulation study was conducted to assess finite sample performances of the tests against sixteen alternatives. A data analysis in the actuarial context is also presented to illustrate application of the testing procedures.

Appendix

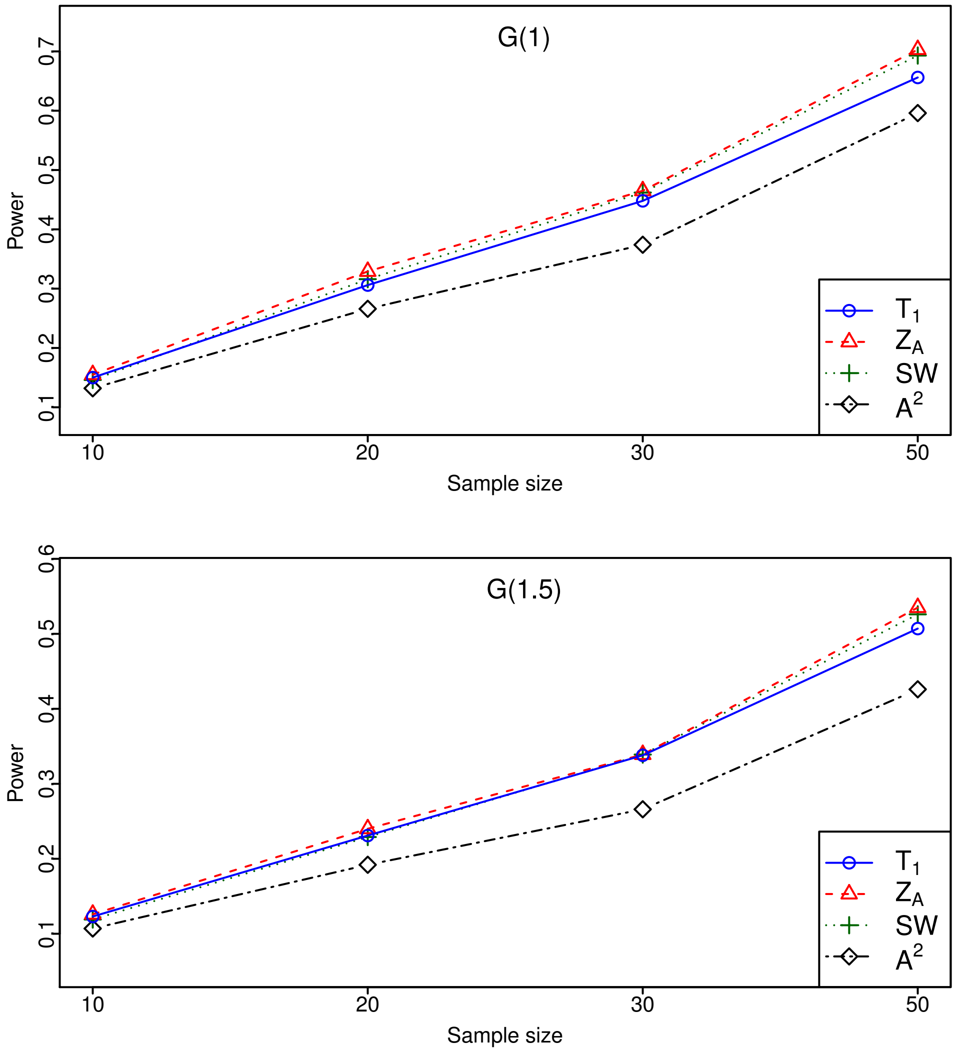


Figure 1: Powers of the tests against G(1) and G(1.5) alternatives.

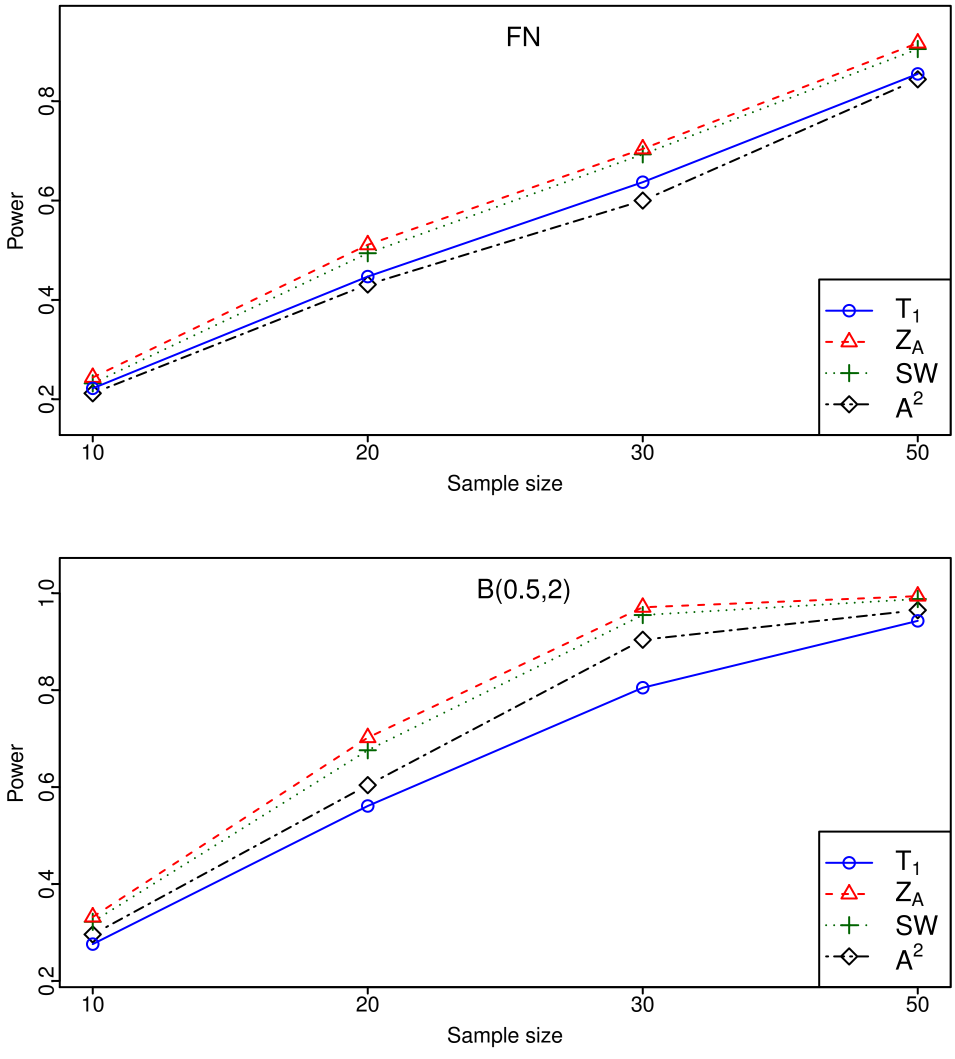


Figure 2: Powers of the tests against FN and B(0.5,2) alternatives.

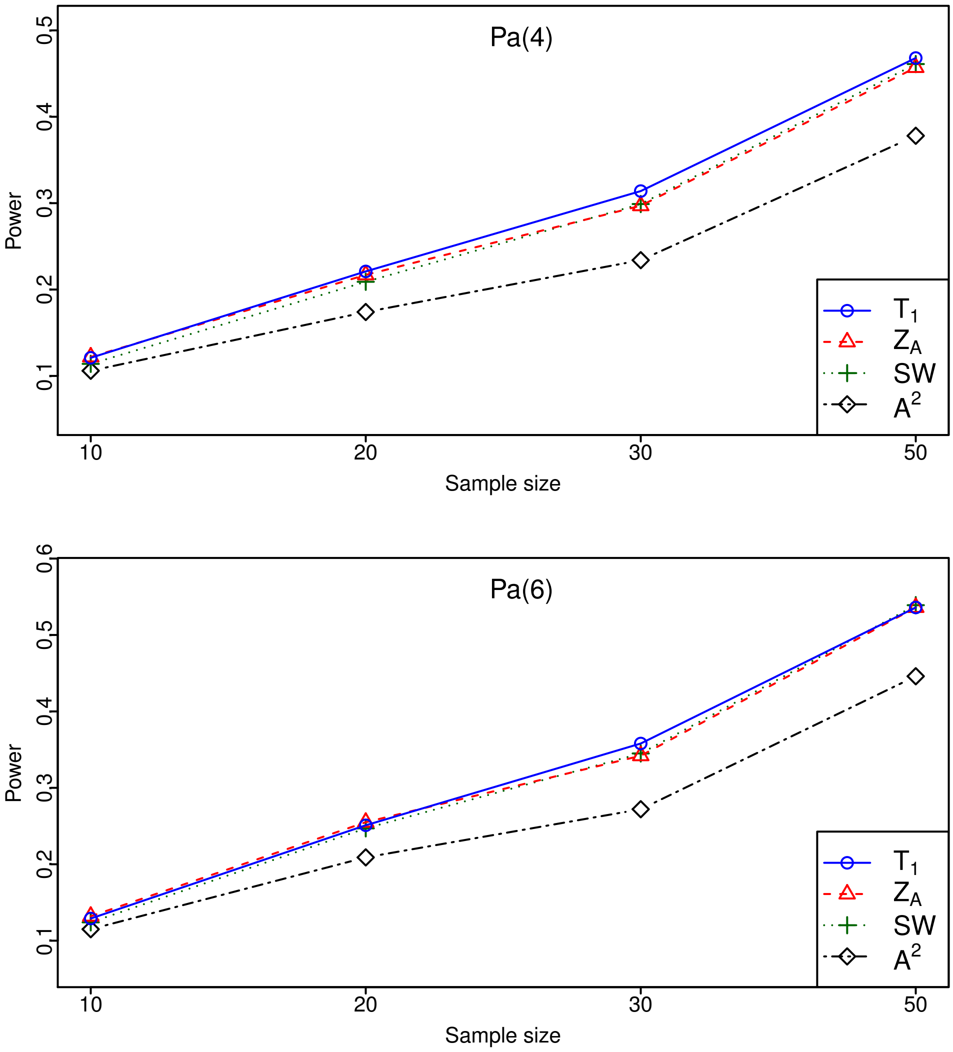
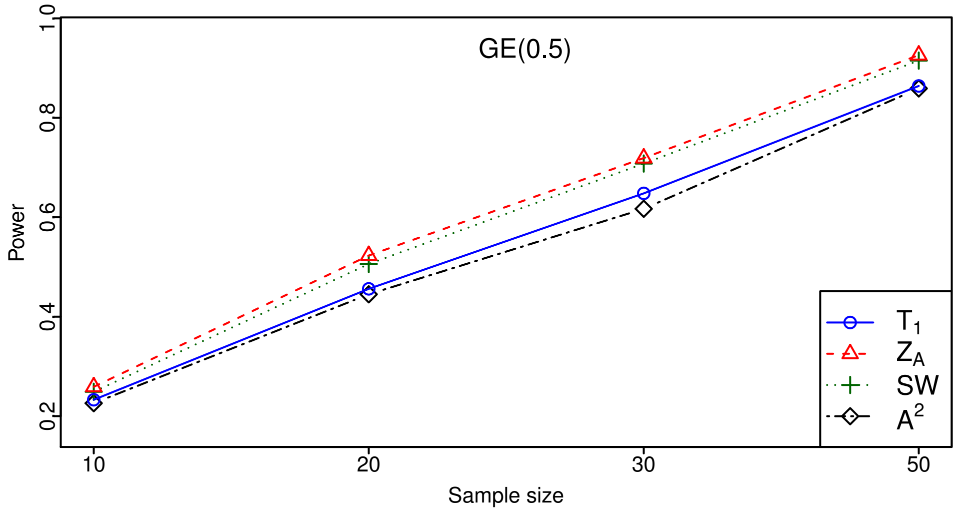


Figure 3: Powers of the tests against Pa(4) and Pa(6) alternatives.



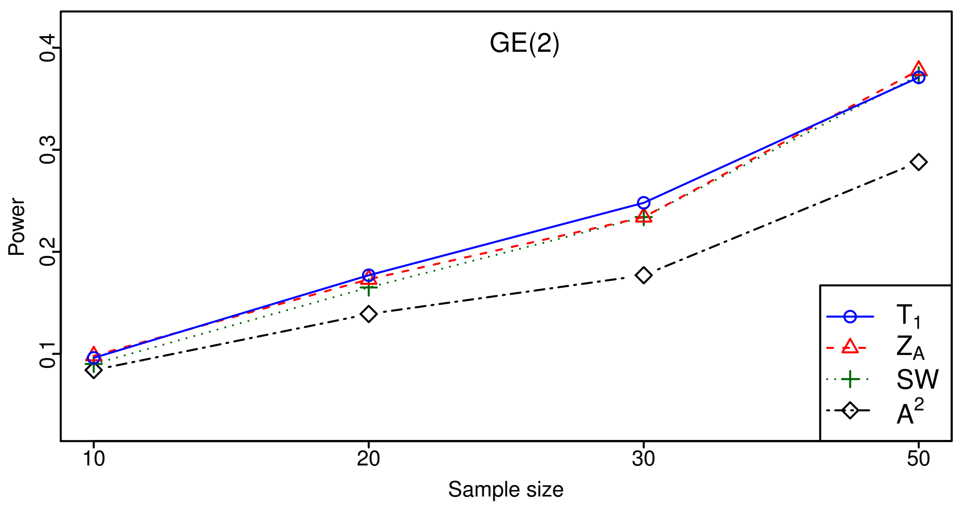


Figure 4: Powers of the tests against GE(0.5) and GE(2) alternatives.

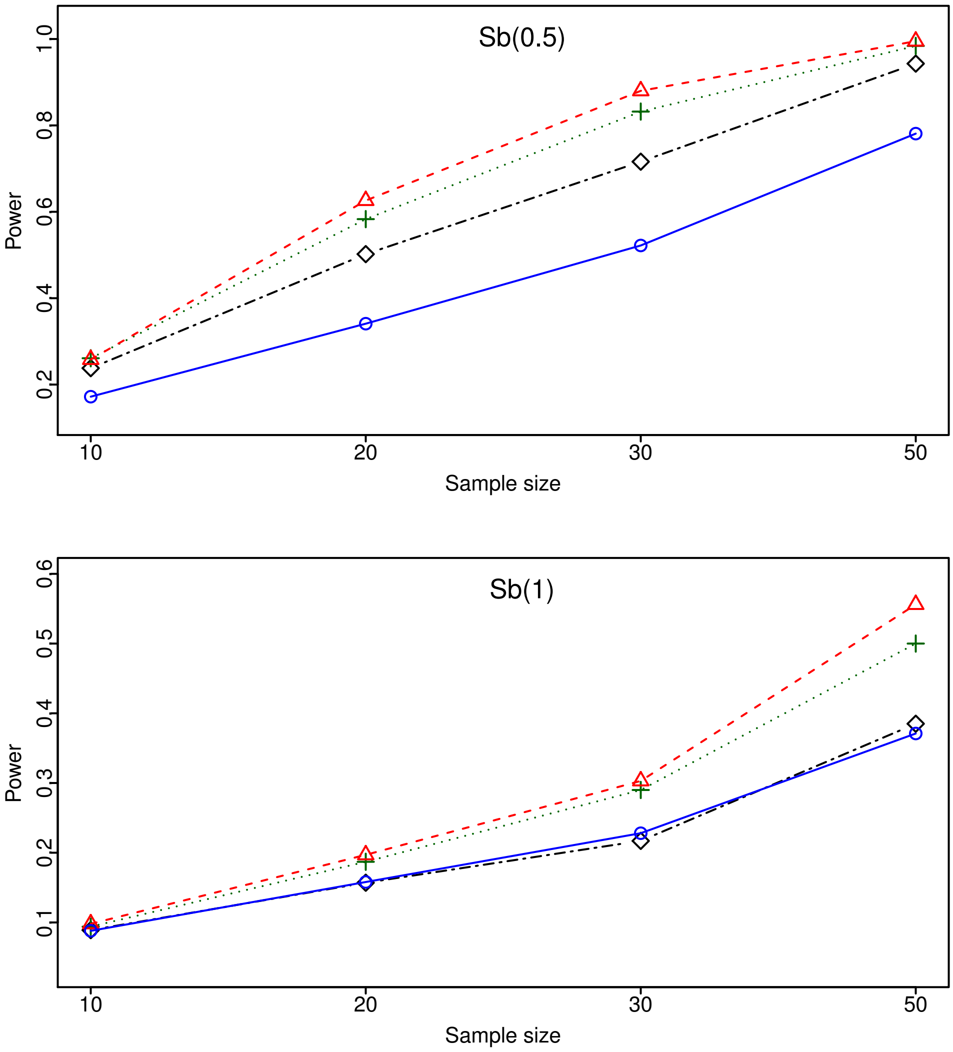


Figure 5: Powers of the tests against Sb(0.5) and Sb(1) alternatives.

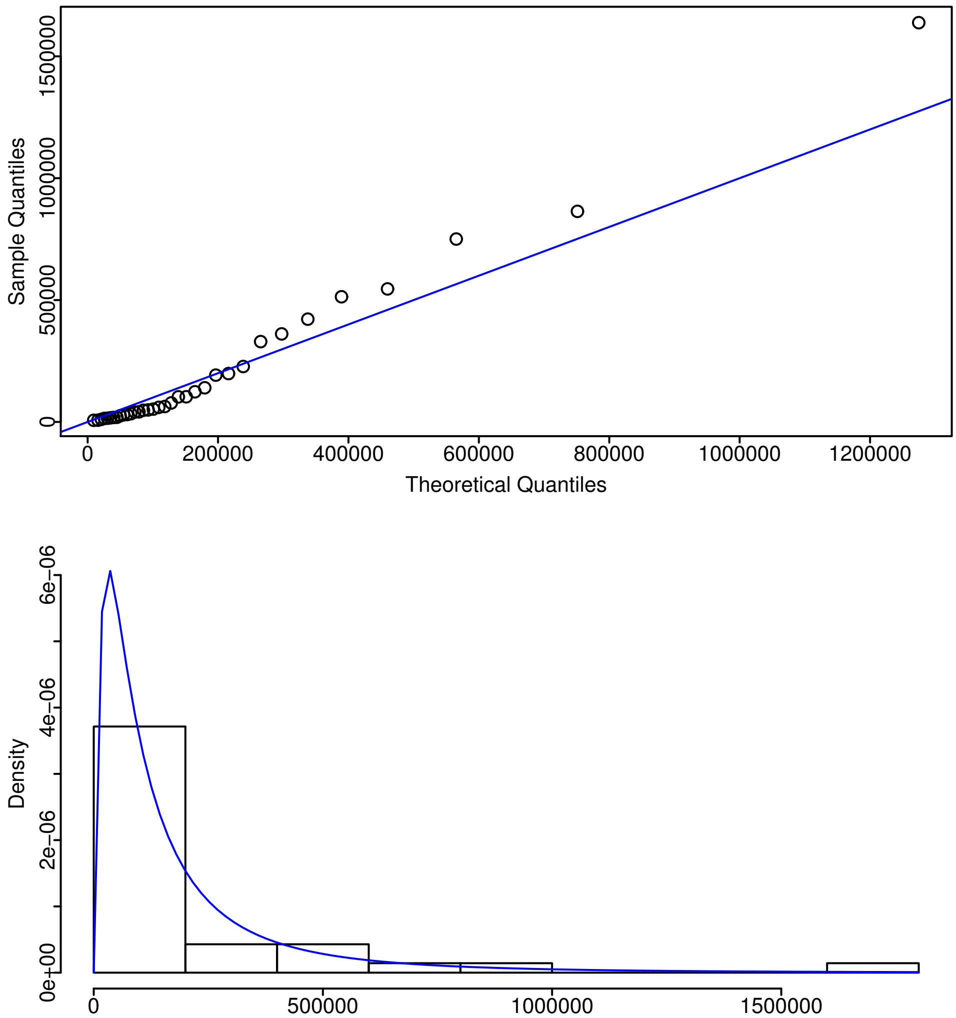


Figure 6: The Q-Q plot, and histogram along with the corresponding fitted LN curve for the hurricane losses data set.

References

Batsidis, A., Economou, P., and Tzavelas, G. (2016). Tests of fit for a lognormal distribution. *Journal of Statistical Computation and Simulation*, 86(2):215–235.

Boland, P. J. (2007). *Statistical and probabilistic methods in actuarial science*. CRC Press.

Bowman, A. (1992). Density based tests for goodness-of-fit. *Journal of Statistical Computation and Simulation*, 40(1-2):1–13.

Correa, J. C. (1995). A new estimator of entropy. *Communications in Statistics-Theory and Methods*, 24(10):2439–2449.

Crow, E. and Shimizu, K. (1988). Lognormal distributions: Theory and practice. *Marcel Decker*.

Gibbons, J. D. and Chakraborti, S. (2011). *Nonparametric statistical inference (5th Edition*. Chapman & Hall/CRC.

Hogg, R. V. and Klugman, S. A. (1983). On the estimation of long tailed skewed distributions with actuarial applications. *Journal of Econometrics*, 23(1):91–102.

Jarque, C. M. and Bera, A. K. (1987). A test for normality of observations and regression residuals. *International Statistical Review*, 55:163–172.

Kirk, R. E. (2008). *Statistics: An Introduction (5th Edition)*. Thomson/Wadsworth.

Koch, A. L. (1966). The logarithm in biology 1. mechanisms generating the log-normal distribution exactly. *Journal of theoretical biology*, 12(2):276–290.

Koch, A. L. (1969). The logarithm in biology: Ii. distributions simulating the log-normal. *Journal of Theoretical Biology*, 23(2):251–268.

Kroll, C. N. and Stedinger, J. R. (1996). Estimation of moments and quantiles using censored data. *Water Resources Research*, 32(4):1005–1012.

Kuiper, N. H. (1960). Indagationes mathematicae (proceedings). 63:38–47.

Magurran, A. E. (1988). *Ecological diversity and its measurement*. Springer Netherlands.

Masuyama, M. (1984). A measure of biochemical individual variability. *Biometrical journal*, 26(3):337–346.

Royston, P. (1993a). A pocket-calculator algorithm for the shapiro-francia test for non-normality: An application to medicine. *Statistics in medicine*, (2):181–184.

Royston, P. (1993b). A toolkit for testing for non-normality in complete and censored samples. *The statistician*, 42:37–43.

Shannon, C. and Weaver, W. (1949). The mathematical theory of communication university of illinois press urbana google scholar.

Shapiro, S. S. and Francia, R. (1972). An approximate analysis of variance test for normality. *Journal of the American Statistical Association*, 67(337):215–216.

Shapiro, S. S. and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4):591–611.

Tzavelas, G. and Economou, P. (2012). Characterization properties of the log-normal distribution obtained with the help of divergence measures. *Statistics & Probability Letters*, 82(10):1837–1840.

Van Es, B. (1992). Estimating functionals related to a density by a class of statistics based on spacings. *Scandinavian Journal of Statistics*, 19:61–72.

Vasicek, O. (1976). A test for normality based on sample entropy. *Journal of the Royal Statistical Society. Series B (Methodological)*, 38:54–59.

Watson, G. S. (1961). Goodness-of-fit tests on a circle. *Biometrika*, 48(1/2):109–114.

Yousefzadeh, F. and Arghami, N. R. (2008). Communications in statistics: Simulation and computation. 37(8):1479–1499.

Zamanzadeh, E. and Arghami, N. (2009). Normality and exponentiality tests based on new entropy estimators. *Journal of Statistical Sciences*, 2(2):179–200.

Zhang, J. and Wu, Y. (2005). Likelihood-ratio tests for normality. *Computational statistics & data analysis*, 49(3):709–721.